



TITLE:

戦後教育評価史に関する覚書

AUTHOR(S):

田中, 耕治

CITATION:

田中, 耕治. 戦後教育評価史に関する覚書. 教育方法の探究 2015, 18: 1-11

ISSUE DATE:

2015-04-14

URL:

<https://doi.org/10.14989/198370>

RIGHT:

戦後教育評価史に関する覚書

田中 耕治

1. はじめに

本章の構成と時期区分

教育方法学研究において、「教育評価」の理論や実践が明確に学的位置を確保するのは、近年のことに属する。事実、日本教育方法学会が編集した『戦後教育方法研究を問い直す—日本教育方法学会30年の成果と課題—』（『教育方法』24, 明治図書, 1995年10月）、同じく本学会編集の『戦後50年、いま学校を問い直す』（『教育方法』25, 明治図書, 1996年10月）においては、とり立てて「教育評価」という項目はない。2009年の本学会編の『日本の授業研究』下巻、学文社において、ようやく「第5章 教育評価と授業研究（西岡加名恵著）」として、明確な位置を与えられている。戦後の教育世界においては、特殊日本の事情もあって、「教育評価」は教育心理学の対象または一領域としてながく定位していたことがそのおもな理由のひとつであろう。

このように学会の公式レベルに即して「教育評価」研究の経緯を概観してみると、戦後史に限っても、その時期区分を厳密に確定することは、今後の課題に属する（なお、1945年から1994年の時期に発刊された教育評価に関する文献一覧は、田中, 1996年資料編また戦後教育評価史研究に関する筆者の基礎作業としては、田中, 2009と田中, 2010aを参照）。

本章では、あくまでも仮説的な時期区分と断ったうえで、日本のようにナショナル・カリキュラム（この場合は「学習指導要領」と「指導要録」の両方）が自国の教育活動の理論面や実践面に与えた大きな影響力を考慮して、戦後指導要録の歴史を次のように仮設してみた（要録史の詳細な分析、検討は田中, 2010b 参照）。本論の題目として「覚書」とした所以である。

第一期 1948年版指導要録

—戦前の「考査」への反省と「指導機能」重視—

第二期 1955年版指導要録, 1961年版指導要録, 1971年版指導要録

—「相対評価」の強化と矛盾の激化—

第三期 1980年版指導要録 1991年版指導要録

—矛盾の「解消」としての「観点別学習状況」欄の登場—

第四期 2001年版指導要録 2010年版指導要録

—「目標に準拠した評価」の全面採用、「目標に準拠した評価」と「個人内評価」の結合

本章では、以上の指導要録の戦後史を念頭において、特に各改訂期を画期とするその理論的展開に着目して、以下のようなプロットを構成してみた。

1. 本章の構成と時期区分

2. 教育評価研究の歴史

(1) 戦後初期「教育評価」概念の成立

(2) 戦後教育評価の多層構造

(3) 「相対評価」批判と「到達度評価」の登場と課題

3. 教育評価研究の課題と展望

(1) PISAと「真正の評価」

(2) 評価研究の新しい動向

2. 教育評価研究の歴史

(1) 戦後初期「教育評価」概念の成立

「教育評価」という用語は、第二次世界大戦後にアメリカから移入された「エバリュエーション (evaluation)」を翻訳したものである。最近復刻された「教員養成のための研究集会記録 (二)」(主催 文部省, 東京帝国大学, 1947年)——この文章は、戦後初期に取り組まれることになった「教育指導者講習

(IFEL)」の前史として知られる研究集会の記録——によれば、そこでは、すでに「エヴァリエーション (evaluation)」という、日本では「教育評価」と邦訳される新しい概念が紹介されている。それも、「エヴァリエーション」概念の提唱者であるタイラー (Tyler, R.W.) の「八年研究 (Eight-Year Study)」にまで目配りするなど、当時の教育改革に対する熱気的一端を伝えている。

現在の学習指導要領では、残念ながらその記述は消滅しているが、戦後初発の「学習指導要領 (1947 年一般編試案)」では、「第五章 学習結果の考査」において、さらには 1951 年の「学習指導要領一般編試案」では、「IV 教育課程の評価」「V 学習指導法と学習成果の評価」というように「教育評価」の項目が明確に位置付けられていた。

そこでは「教育課程の評価と教育課程の改善とは連続した一つの仕事であってこれを切り離して考えることはできない。この意味において、教育課程の評価は、教育課程の計画、その展開とともに、児童・生徒の学習を効果的に進めていく上に欠くことのできない仕事である。」とされ、さらには「その意味で、自ら実施した活動について、絶えずあらゆる機会においてそれを検討し、評価し、これに改善を加えていく責任が、とりわけ個々の教師には課せられている。」(1951 年の学習指導要領一般編試案より)と明記されており、まさしく、戦後初期の学習指導要領が「試案」とされたことと整合した記述となっている。

また、戦後最初の「指導要録 (当初は『学籍簿』と呼称、1949 年に「指導要録」と改称) では、戦前の「学籍簿」のあり方を反省して、文部省学校教育長名の通達「小学校学籍簿について」(1948 年 11 月 12 日)において、「小学校学籍簿の趣旨」として、次の三点が述べられている。「1. 個々の児童について、全体的に、継続的に、その発達の経過を記録し、その指導上必要な原簿となるものである。2. 記録事項は、新しい教育の精神からみて、とくに重要と思われるものを選定してある。3. 出来るだけ客観的に、しかも簡単に、かつ容易に記録されるように作られてある。」とされ、とりわけその 1 において「指導機能」が重視されていたことが理解できるだろう。このような文部省の方針を受けて、戦後初期において「教育評価」に関する豊

かな実践が展開されたことが、近年の研究において発掘されている(林, 1997 や坂井, 2009 参照)。

上述の当時の文部省の「教育評価」理解の水準を示す典型的な事例として、文部省が作成した『初等教育の原理』(1951 年)をあげることができよう。

そこでは、「教育評価」概念の特徴が次の五点にわたって明快に説明されている。①評価は、児童の生活全体を問題にし、その発展をはかろうとするものである。②評価は、教育の結果ばかりでなく、その過程を重視するものである。③評価は、教師の行う評価ばかりでなく児童の自己評価をも大事なものと取り上げる。④評価は、その結果をいっそう適切な教材の選択や、学習指導法の改善に利用し役だてるためにも行われる。⑤評価は、学習活動を有効ならしめる上に欠くべからざるものである(文部省, 1951, pp.217-219)。

すなわち、「教育評価」とは、教育活動を評価することであって、教師の指導と子どもたちの学習活動(自己評価を含む)の改善をめざす行為であると明快に規定されている。

日本に導入された「エヴァリエーション」概念は、1920 年代に隆盛をきわめた教育測定運動が提起した「測定」概念を否定的に媒介することによって、1930 年代にタイラー (Tyler, R.W., 1934) によって提起されたものである。タイラーは、「教育評価」においては教育目標こそ妥当性の本質的な基準であり、あくまでも測定の「信頼性」に固執して、統計的尺度法の開発を優先する教育測定運動の形骸化を批判した。

当時の日本の教育心理学者の中で、「測定」概念を批判して、タイラーの「教育評価」の意義を最も深く理解したのは正木正 (1905-1959) であった。

正木は、測定や評価行為を「教育現実の構造や歴史的展開の道」(正木, 1952, p.246)に即して考察することで、この問題を「技術の末梢的議論のうちに終止」(正木, 1952, p.244)する傾向を批判する。そして、「教育的価値・目標」を問わないと、評価の管理的・測定的性格は「人間疎外」に陥ると警告する。たとえば、「測定基準」とは帰納的経験的に求められる「統計的平均」を意味するが、「評価基準」は、「郷土社会のもつ歴史的社会的規定に立つ規準」(正木, 1948, p.352)であるとして、その相違を明確にしている。このようにして正木は、まさしく今日に通じる教育(方

法)学における「教育評価」の意義を次のように析出している。「今日周知の如く教育過程において評価(evaluation)が重要視されてきている。これは評価という問題が教育過程に新しく付加され、関心されて来たということではない。むしろ、評価は教育過程に融合されている部分であって、従来問題的にも、方法的にも無意識であったものが、新しく自覚されて来たということである。教育活動の自己発展として評価の問題が分節化し、強調されて来たというべきであろう」(正木, 1952, p.245)。以上、正木の提起は、戦後初期の「教育評価」概念の理念・目的を真に実現していく際の方法論上の留意点を指摘したものであった。しかしながら、これらの提起は、当時においては必ずしも主流となりえなかった。この提起は、後年續有恒の所論に継承されていく。

(2) 戦後教育評価の多層構造

今日の時点の教育評価研究の到達点から俯瞰すれば、「教育評価」においてタイラーは教育目標の拡張(高次の認知機能や情意的側面を含む教育目標の設定)を図ることによって、教育実践との妥当性を追究したが、その拡張された教育目標を評価する方法、さらにはその評価結果をいかなる評価基準に依ってかかるのかについては十分に詰めた議論がなされておらず、その結果として評価の信頼性、客観性を重視する「測定」概念が温存されることになったと考えてよいだろう(当時、アメリカにおける「教育評価」論争を紹介した長島貞夫, 1949 を参照されたい①)。もちろん、それ以上に、学歴社会に突入しようとしていた当時の日本においては、ハイスティクスな性格を帯び始めた「内申書」の原簿となる「指導要録」に「証明機能」が強く求められ、「5段階相対評価」が1955年改訂の指導要録において、総合評定(戦後最初の指導要録で採用された「相対評価」は分析評定であった)として採用され、いわば「相対評価」があたかも「教育評価」を独占して、跋扈し始めたのである。その後の日本における教育評価研究は、この「相対評価」をいかに批判し、克服するのかについて焦点化されていくことになる。

ところで、「相対評価」の独占と記述したが、当時の指導要録の構造はそれほどモノトーンではない。戦後の教育評価研究をリードした橋本重治

(1908-1992)によれば(橋本, 1961, pp.356-362)、それぞれの教育評価観の長所と短所をあげることで、とくに「絶対評価(この場合は現代の「目標に準拠した評価」)」に傾斜しようとする動向を牽制し、「相対評価」を擁護することである。筆者は、その影響力の大きさから、それを橋本パラダイムと呼んでおきたい。

橋本パラダイムによれば、まず、「相対評価」は「教育学」的にみた場合多くの欠点があることを承認する。そこで、「絶対評価」が着目されるが、それは、「教育学」的にはすぐれているが、「教育測定学」的にみた場合には主観的で信頼性に乏しい。したがって、外部証明のための機能を持つ指導要録では不適格な評価論である。この目的のためには、やはり、「教育測定学」的に「客観性」「信頼性」をもつ「相対評価」が有効である。しかし、その「教育学」的な欠点があるために、それを補うために、「個人内評価」が必要とされる。これが、橋本の論法である。そして、この考え方を背景として、1961年改訂の指導要録では、「評定」欄では相対評価、「所見」欄では横断的な個人内評価、「進歩の状況」欄では縦断的な個人内評価というように、配置されていると説明される。ただし、この説明では絶対評価の位置が明確でないが、それは、「平常の1回1回のテスト」で採用されているし、さらには、「評定」欄において、「相対評価に絶対評価を加味する」という形で位置づくことされる(ここで挙げられている「加味」される「絶対評価」とは事実上「個人内評価」のことである)。以上のことから、戦後半世紀にわたる公的な教育評価の構造は、「相対評価」を基軸とする多層構造であったと言ってよいだろう。

さて、上述の橋本の論法において、「個人内評価」の役割が注目されていることに着目しておきたい。戦後の教育評価論を語る場合に忘れてはならないことは、「相対評価」とともに「個人内評価」が、戦後最初の指導要録から採用されたことである。この「個人内評価」は、子どもたちの目線に合わせて、子どもたちを発達的に全体的に評価しようとするものであって、明らかに教師を絶対視する戦前の「考査」を反省して登場してきたものである。

ところが、この「相対評価」と「個人内評価」の関係が、教育現場においては、社会学が指摘する「過熱と冷却の構造」をなしており、さらには『現代教育科

学』(1961年3月)に掲載された「相対評価」に対する批判の二つの系譜〔遠山啓らの『教育と科学の結合』論と東井義雄らの『教育と生活の結合』論〕を析出して、筆者は「到達度評価と個人内評価の結合」を提唱した(田中, 1996)。

すなわち、「相対評価」批判の二つの系譜とは、そのひとつは、遠山啓(1909-1979)や桑原作次の立場である。遠山啓は、成績を上げるためには誰かが落ちなくてはならないから、相対評価は排他的な競争を人為的に煽り、「他人の不幸はおのれの幸福」(p.4)とする心情を形成することになると批判した上で、「もともと教育の目標はすべての子どもが5をとることであり、教師の努力もそこを旨ざしているはずである。ところが、五段階評価はそれを真つこうから否定していることになる。」(p.4)と述べている。また、桑原作次は、教育という営みは、自然状態にある大数の傾向である正規分布曲線を崩して、「教育曲線」とでもいうべきものに作り替える作業であるから、「教育の結果、依然として正常曲線型を示すとすれば、それは、問題が適切であったなどと喜ぶよりも、教育そのものが適切であったかどうかについて、むしろ教師の反省を要求しているものとみるべきではないか」(p.33)と明快に述べている。これらの「相対評価」批判は、後の「到達度評価」の登場につながる、すべての子どもたちの学力保障を願う立場から発せられたものである。

他方、同誌上には東井義雄(1912-1991)や坂本光男のように、おもに生活綴方の伝統に接続する立場から、「相対評価」の非教育性を抽象化された点数化にあるとして、子どもたちの内面の成長・発達を記述する方法原理として「個人内評価」を押し出そうとしている。東井義雄は、「一学期の『3』よりは二学期の『3』には質の高まりがあるはずだ。このちがいを、同じ『3』の中に見よという方が無理ではないか。」(p.10)と述べて、「私は、成績を数字であらわすことには、相対評価であろうと絶対評価であろうと反対したい。」(p.11)と主張している。ちなみに、大正自由教育の系譜に属する斎藤喜博(1911-1981)もまた、戦前においてすでに「学級全体の横の関係」ではなく、「個人の縦の進歩」こそ重視すべきことを指摘していた。

(3)「相対評価」批判と「到達度評価」の登場と課題

1969年2月に起こった「通信簿事件」は(田中, 2008a, p.177), テレビのワイド番組で取り上げられたこともあって、「相対評価」の不条理を国民レベルで顕在化させ、以後全国の小学校現場で通知表改革が展開され、1971年の指導要録改訂にも影響を及ぼし「5段階に機械的に割り振ることのないように」と注記され、「相対評価」の後退がみられるようになる。その際、この通知表改革が単なる名称変更や5段階を機械的でない配分率にするといった、結果として様式イジリに陥ろうとしていた時に、「方向目標＝相対評価、到達目標＝到達度評価」というシエマ(中内, 1977)を念頭に置いて、「到達度評価」論を樹立するのが中内敏夫である。中内は日本の教育学者としておそらく最初に「教育評価」の今日的意義と役割を体系的に提起した研究者であり、『学力と評価の理論』(1971年, 国土社)はその金字塔となる著作である。

ところで、「到達度評価」という名称は、京都府教育委員会作成の『到達度評価への改善を進めるために―研究討議のための資料―』(1975年2月)いわゆる「長帳」において、初めて登場した。そこには、次のように書かれている。

「到達度評価は、すべての子どもの学力を、それぞれの学年、教科の目標に到達させることを基本とする教育指導における評価なので、すべての子どもの以後の発達にとって、必要で、十分な教育を整えていくことに努める教育指導の全体に役立つものとしなければなりません。」(p.5)

この「到達度評価」の成立(佐々木, 1979)に、直接間接に与った理論や研究者は、教育心理学者の續有恒(續, 1973), 村越邦男(村越, 1978), ブルーム(Bloom, J.S.)の理論を紹介、翻案した梶田叡一(梶田, 1979), 京都の到達度評価論の理論的支柱である稲葉宏雄(稲葉, 1984)がおり、もちろんその主要なトレーガーとして学習権保障論から創発した中内敏夫がいた。同じく「到達度評価」を主張した論者たちの異同を明確することも今後の課題であろう。

ちなみに、正木正の教え子である續は、従来の「履修主義」を批判して、「修得主義」を主張して、「教育評価」を次のように規定している。「評価は、目標追求

一評価一調整という単位での、目標追求活動における部分活動であって、追求活動の実績と目標との関係をチェックし、調整活動のために、フィードバック情報を提供するものであること」(續, 1969, p.27)。明快な規定であり、その「修得主義」の提唱と合わせて、續の主張は学力保障論を主張する「到達度評価」と接続させるものであった。ただし、先述した『初等教育の原理』(1951年)と比較対照してみると、とくに②評価は、教育の結果ばかりでなく、その過程を重視するものである。③評価は、教師の行う評価ばかりでなく児童の自己評価をも大事なものと取り上げる。という観点が後退し、續の規定にはやや機械的でシステマチックな印象が残る。ここには、当時紹介されつつあった行動主義的なプログラム学習理論の影響または残滓をみることができ、後年「到達度評価」批判を呼び込む結果となったと考えてよいだろう。

ここでは、「到達度評価」によって、歴史的に蓄積されてきた「5段階相対評価」の批判点について整理してみよう。

まず、必ずできない子がいるということを前提とする非教育的な評価論であることである。たとえば、40人学級では「5」は「2~3人」つけることになるが、一方で「1」も「2~3人」つけなくてはならない。ここには、良くできる子どももいるが、いかに指導しようとも必ずできない子どもが存在するという、1920年代に隆盛をきわめた教育測定運動を支えた「社会ダーウィニズム(優勝劣敗)」の思想が根底に流れている。東井義雄)をはじめとする心ある教師たちが「5段階相対評価」の通知表をつける時に戸惑ったのはこの点であって、指導の甲斐あってクラスみんなが頑張って勉強したのに「2」や「1」を無理矢理にも割り振らなくてはならない不条理に心痛めたのである。

この点をたとえば東井義雄(1912-1991)によれば、「学期末を迎えるたびに、私は長嘆息する。手にもつペンは遅々して進まない。遂には目をつむって『2』と書き込む。そしてまた、ため息をつく。・・・“あんなにがんばっていたのに、通知表に記入しなければならぬのは、やっぱり『2』だ。」「自分を表現する力ができたぞーと賞め続けてきた遅進児のT君。でもやっぱり、通知表には『1』だ」・・・終業式の後、通知

表を渡し終えた私は、いつだって頼のこわばるのを感じる」(東井, 初出 1961, 1972 所収, p.7)と述べている。そして、東井は、八鹿小学校の校長時代に通知表改革に乗り出す(東井, 1967)。

「相対評価」の第二の問題点は排他的な競争を常態化させて、「勉強とは勝ち負け」とする学習観を生み出すことである。「評点(語)」の配分率があらかじめ決められている場合には、成績を上げるためには誰かが落ちなくてはならないことから、「相対評価」は排他的な競争を人為的に煽り、「他人の不幸はおのれの幸福」と(遠山, 1961, p.4)する心情を形成することになる。もちろんここでは、教育における競争一般(たとえば、全員5をとるための競争など)が否定されているのではなく、まさしく「排他的な」競争が問題視されているのである。

「相対評価」の第三の問題点は、学力の実態を映し出す評価ではないことである。「5段階相対評価」でたとえ「5」をもらったとしても、その意味するところは集団における相対的な位置が上位であるということであって、そこで獲得した学力がめざすべき教育目標に達していたかどうかを証明してはいない。むしろ、「勉強とは勝ち負け」であるという学習観が強化されることと反比例して、「何を勉強したのか」という問いは確実に希薄化していくことになる。「受験学力」の大きな問題点のひとつはここにあるといつてよいであろう。

第四の問題点は、「相対評価」が子どもたちの学習活動をネグミして、集団の中での序列化を行うことを目的とする評価法であるかぎり、そこでは子どもたちの努力(不足)や能力(不足)のみが問題とされることになる。したがって、本来の教育評価の目的や役割である、「評価結果は教育活動の反映であり、それに基づいて教師はみずからの教育力量や取り巻く教育環境を点検し、かつ鍛えるという回路」を持たないことになる。「勉強とは勝ち負け」とする学習観が跋扈するのに反して、学力を形成する教育活動の空洞化が確実に進行することになるのである。

「相対評価」を以上のように批判する「目標に準拠した評価」である「到達度評価」は、以下のように主張する。

「到達度評価」の何よりの意義は、本来の教育評価

の役割を復権させたことである。既述したように、「教育評価」の役割とは、すべての子どもたちの学力保障をめざして、まさしく教師たちの指導のあり方を点検し、子どもたちの学習活動に見通しを与えるために機能しなくてはならないものである。繰り返すまでもなく、「相対評価」では、子どもたちの能力や努力は問われることがあっても、教育実践のあり方を問い直すという契機はない。それに対して、「目標に準拠した評価」は目標が子どもたちに実現していく度合いを捉えて、もし不十分な場合には、教師の指導を修正したり、子どもたちの学習活動を支援していくことになる。

さらに、「到達度評価」は本来の教育評価の役割を復権させようとするものであって、その実現をめざして「評価機能の分化」を提起した。学力の保障をめざす「到達度評価」では、授業に入る前の子どもたちの学力実態を把握する「診断的評価」、授業中の子どもたちのさまざまな反応におうじて指導の内容や方法を調整する「形成的評価」、そして目標がどの程度実現したのかを捉える「総括的評価」と、三つの評価によって子どもたちの学力形成の実相をきめ細かに捉えようとした。とりわけ、「形成的評価」は指導に生かす評価の典型として、その後市民権を得ていくことになった。この評価機能の分化については、当時紹介されたブルーム（ブルーム他、1973、1986 参照）の影響がある。また、ブルームがタイラーの着想を継承した『教育目標の分類学（タキソノミー）』は、教育目標の内容面と同時に能力（行動）面を階層化する提案であって、全米にとどまらず世界的な影響力を持つものであった（その改訂作業については、石井、2010 参照）。

さて、日本では「相対評価」批判が顕在化して「目標に準拠した評価」としての「到達度評価」が強調されようとしていた 1970 年代中頃におきた、その後の教育評価研究にとって無視できない事項に触れてみたい。この時期に、文部省と経済協力開発機構（OECD）の教育研究革新センター（CERI）との共催による「カリキュラム開発に関する国際セミナー」（1974 年に東京で開催）において、アトキン（Atkin, J.M.）がカリキュラムと授業さらには評価に関わる典型的なふたつのモデル（「工学的アプローチ」と「羅生門的アプローチ」）を提案した。

「工学的アプローチ（technological approach）」とは、

まさしく「工学」に例えられるように、教師の意図的な計画化とそれに基づく「行動目標」と教材配列による授業の合理的な組織化の様相をあらわしている。そして、そこで実施される評価論は、「目標に準拠した評価」である。

他方、「羅生門的アプローチ（rashomon approach）」は、子どもたちの能動的で多面的な学習活動を展開するために、一般的な目標のもとに創造的で「即興を重視する」授業が展開される。そして、そこで実施される評価論は、「目標にとらわれない評価」（平野、1981 参照）である。子どもと教師と教材との「出会い」から生まれる学習の価値をさまざまな立場や視点から解釈の様相が、芥川龍之介の小説『藪の中』を黒沢明監督が映画化した『羅生門』（1950 年）と重なることから、このように命名された（文部省、1975）。

このように二つの考え方を対比することは、「工学的アプローチ」と「目標に準拠した評価」を相対化する試みである（田中、1982 参照）とともに、その根底には「工学的アプローチ」と「目標に準拠した評価」に対する明らかな批判意識が存在していると考えてよいであろう。ここでは、「目標に準拠した評価」としての「到達度評価」に対する理論的批判を以下のようにまとめておきたい。ここでは、既述したように、續の評価規定が、プログラム学習理論の影響または残滓によるややシスタマチックな印象を与えると指摘したことを想起していただきたい。

「目標に準拠した評価」の典型的な形態であった「到達度評価」に対しては、はやくから教師の構想した教育目標に子どもたちが到達しているか否かを点検しているのに過ぎないのではないかという疑問や批判が寄せられていた。それは、「目標つぶし」「詰め込み主義」の授業を助長するのではないかという懸念でもあった。急いで断っておくと、「到達度評価」として実践された授業のすべてにこのような問題点が発生していたわけではない。また、そのような疑問や懸念を克服するための努力も積み重ねられてきたことも事実である。アメリカにおいても、ブルームの教え子たちによって、マスタリー・ラーニングとプログラム学習の相違が追究されていた（田中、1995 参照）。

ここでは、まずは「目標に準拠した評価」をめぐるてなされた批判を四点にわたって整理してみたい。こ

の作業によって、「目標に準拠した評価」の課題として「個人内評価」との結合、さらには構成主義的学習観との結合が求められようになったことが理解できるであろう。

まずは、「目標に準拠した評価」は、教師による「目標」が規準となることから、それからはみ出す子どもたちの活動を見落とす危険性があるのではないか。それを防ぐためには、まずは子どもたちの、ひいては保護者や地域住民の教育評価への「参加」が保証される必要があり、そのことによって、多面的に多層的に教育活動を捉えることが可能になってくるのではないかという批判であった。これらは、スクリヴアン (Scriven, M.) による「ゴール・フリー (目標にとらわれない)」評価 (根津, 2006 参照) からの批判と総称されることになる。

この「目標に準拠した評価」に対して、早くから批判的見解を述べていたのは、芸術教育に関わっていたアイスナー (Eisner, W.) であり、同じく「羅生門的アプローチ」の論拠となったものである (Eisner, 1985) (桂, 2009 参照)。

アイスナーの批判は、次の言葉に凝縮されている。「教授の動的で複雑な過程は余りにも多数の成果を生むので事前に行動的で内容的な用語で特定化することはできない」 (Eisner, 1985, p.254) 「教育目標をそれによって教育的達成を測定する規準とみなすカリキュラム理論は、測定不可能な達成の諸様相を見落とす」 (Eisner, 1985, p.257)

すなわち、「目標に準拠した評価」は、教授学習のプロセスをまるで工場の組立工程のようにみなしているのであって、評価行為においては、基準の適応 (the application of a standard) と判断の行使 (the making of a judgement) を区別すべきであると主張する。

次に、同じく「目標に準拠した評価」は、子どもたちにとっては、「外的な評価」であって、子どもたち自身による「内的な評価」を充分には位置づけていないのではないか。すると、子どもたちは自分の行う活動の値打ちづけを常に他者からの評価に依存することが習性になって、いわゆる「指示待ち」「評価待ち」の状態になってしまう。「目標に準拠した評価」において、子どもたちの自己評価をどのように活性化させるかが課題となってくるのではないかという批判であった。

第三点目として、「目標に準拠した評価」は子どもたちの学習の成果に着目するあまり、その成果や結果に至る「プロセス」を丁寧に読み取ることに課題が残されているのではないか。教師によってもたらされる「知識」が子どもたちの既存の知識体系と掛け離れていれば、そこに「葛藤」や「矛盾」が生じることになる。そして、この「葛藤」や「矛盾」を大切にすることで授業のダイナミズムが保証される。まさに、この「プロセス」の評価への方法論が問われているのではないかという批判であった。

最後に、「目標に準拠した評価」では客観性を重んじるために、客観テストが多用される場合が多い。もちろん、客観テストといえども一様ではないが、それでも測定の範囲はおのずから限定されてしまう。いわゆる量的な評価にとらわれて、高度な学力を対象とする質的な評価やたとえば教室に浸透する雰囲気などの質的な側面への目配りが弱くなる。ここでは、評価課題をもっとリアルなものにするとか、子どもたちの獲得した学力をさまざまな方法で表現させるという工夫が求められているのではないかという批判であった。

つまりは、「質」の高い目標の設定と評価関係者 (ステイクホルダー) の主体的な「参加」の保障を求めようとするものであった。

3. 教育評価研究の課題と展望

(1) PISA と「真正の評価」

周知のように、OECD が実施した PISA が与えた教育評価を含む教育課程改革は、世界規模にわたり、まさしく教育のグローバル化を象徴するものとなっている (佐藤学, 2009, 久田, 2013 など参照)。日本においては、PISA 型学力=活用する力をいかに育成するかについては、すでに学習指導要領の改訂を見通した議論が始まろうとしている (勝野頼彦, 2013)。また、PISA で高得点を維持する中国や韓国の教育改革が注目されるようになっている (項, 2013, 京都大学大学院教育学研究科, 2012 など参照)。旧い「目標に準拠した評価」である「ドメイン準拠評価に代わって、質の高い学力をめざす「スタンダード準拠評価」 (鈴木, 2013) の推奨と各国のスタンダード (とくにオーストラリアの取り組み) が紹介されている。その際には、とりわけ、PISA 型学力 vs 教養型学力②または「真正

の評価」論が提起する「参加」の保障の有無が争点となろうとしている。

教育評価研究の立場からは、PISA がまさしく教育調査、教育評価であることから、その基底にある教育評価論としての「真正の評価(authentic assessment)」論に注目が集まっている。ところで、「真正の評価」論が PISA によって関心がもたれるよりも少し前に、日本では「ポートフォリオ評価」の中で注目されていた。それは、総合的な学習が提起されて、それにふさわしい教育評価の方法として、1999 年以降精力的に紹介・考察された(西岡, 2003)。また、「ポートフォリオ評価」を現場で使いやすくする工夫として、堀哲夫は『一枚ポートフォリオ』を提唱している(堀, 2013)

PISA が部分的に採用した「パフォーマンス評価」と「ポートフォリオ評価」の基底にある「真正の評価」論への高まる関心を念頭に置いて、そのオリジナルな原典の翻訳も精力的になされている(ウィギンズ, マクタイ, 2012. ダイアン・ハート, 2012)。筆者は、「真正の評価」論のメルクマールとして、「質」と「参加」の保障という二つの原理を析出している(田中, 第 1 章, 2013)。とりわけ、その「真正性」について以下のように問題提起した。「筆者はかつて「真正性」には、「親密さ」と「困難さ」という二つのアンビバレントな意味が内包されていることに注意を促した。(田中, 2008 年, pp.73-74) この点について、教育分野にはじめて「真正性」概念を導入したフレッド・ニューマン(Newmann, F.M., 藤本, 2013 参照)たちは、その基本文書の中で、「真正性」を規定する際に、「親密さまたは適切性(Relevance)」と「困難さまたは厳密性(Rigor)」という用語を採用し、元来前者には経験主義的な子ども中心のニュアンスが強く、現実世界で活躍する文化的知的職業に内包されている営みでは後者こそ本質的な側面であると説明している(田中, 2013, p.22)。この「真正性」概念に関する追究は今後の課題であろう。なお、2010 年の指導要録の改訂において、「パフォーマンス評価」が推奨されている。その評価の「信頼性」を確保するために、いかに「ルーブリック」を作成するのが課題となろうとしている(戦後初期には、小宮山栄一は、今日のルーブリックの原初的形態に当たる「記述尺度法」を提案していた(小宮山, 1950, p.79)(田中編, 2001 参照)。

(2) 評価研究の新しい動向

PISA のような大規模な教育調査や教育評価が席卷するようになると、その教育評価はある種のシステマチックとなり、日々の教育実践から乖離する危惧が生じる。鹿毛が指摘するように、「教育評価を実践として豊かなものにしていくためには、ひいては教育の『質』を高めていくためには、何よりもまず、教師たちが『見る目』を磨き、『聴く耳』を鍛えることが必要なのではないだろうか。」(鹿毛, 2007, p.101)と問いかけている。

最近の『教育方法学研究』において、特に若い教育評価研究者たちによって、「学習のための評価」または「学習としての評価」と総称してもよい、教育実践に内在する評価行為の研究が取り組まれている。たとえば、最近のイギリスで取り組まようとしている構成主義的な学習観に基づく「形成的評価」の展開を紹介した二宮論文(二宮, 2012) また、グーバ(Guba, E.G)とリンカーン(Lincoln, Y.S.)によって「第四世代評価」とされる自然主義的アプローチに基づく作文評価の研究を展開した北川論文(北川, 2006)、ヴィゴツキー(Vygotsky, L.S.)理論に淵源を持つダイナミック・アセスメントを算数授業に試みた平田論文(平田, 2007)などに注目しておきたい。筆者は、日本の教師文化の独自性である『教育実践記録』に着目して、「教育評価論としての『実践記録』(田中, 第 7 章, 2013)を考察してみた。「真正の評価」論と実践に内在する評価論との結合をどのように構想するのかも今後の課題である。

さて、教育評価が日々の教育実践に強力に作用するのは、とりわけその制度設計を通してである。とりわけ、日本においては入学試験制度が大きな影響力を持っている。新聞報道(『朝日新聞』2013/11/1, 15/1/9)によれば、政府の教育再生会議を受けた中教審が大学入試改革を答申したと伝えている。それによると、従来のセンター試験に代わる「大学入学希望者学力評価テスト」が構想されている。紙幅の関係で、この分野についてのこれ以上の考察ができないが、この入試改革は今後の日本の教育設計全体に及ぶものとなると考えられよう(田中, 第 6 章, 2008, 荒井, 2005 参照)。稿を改めて論じたい。

参考注

①『教育評価』論争とは, Journal of Educational Research.誌上で争われたものである。

<測定派>

Orata,P.T., Evaluating Evaluation. Journal of Educational Research. May 1940.

Woody,C., Nature of Evaluation. Journal of Educational Research. March 1942.

↑

<評価派>

Tyler,R.W., General Statement on Evaluation. Journal of Educational Research. March 1942.

この論争における測定派の評価派への批判は長嶋は以下のとおりまとめている。

- 一 「これらの評価の用具は主として、測定用具と同様に、ペーパーテストであって、それらに見られる共通の限界からのがれることができないこと。
- 二 測定用具は、妥当性を欠く点で非難されて来たが、評価用具も、児童の成長・発達過程に対して十分の妥当性ありとは必ずしもいえないこと。
- 三 質的差異に重点をおくことは、一方には主観性を容認する結果になり、いかにして信頼度を確保するかが問題とされること。
- 四 測定においては、質的なものを量的に表現することにおいて攻撃されたが、評価用具も結局は同様であるといわなくてはならないこと。
- 五 教育測定は、決まりきった解答を要求する点で欠点を指摘されたが、評価用具においても必ずしも児童の自由な表現による解答は要求されていないこと。
- 六 評価用具も学習過程についての記述よりは結局、学習結果を重視していることなどが挙げられている。

長嶋貞夫「エヴァリュエーション」『新教育事典』平凡社、1949年、pp.27-28。

以上の論争から、質的側面を重視する当時の「教育評価」論に対して、評価基準の信頼性、妥当性と評価方法の多様性が求められていることが理解できよう。「真正の評価」論は、その課題への一つの回答であっ

たと考えてよいだろう。

②PISA のリテラシー論に対しては、Bildung の伝統を持つドイツやエスプリの伝統を持つフランスでは批判的論議が起こっている。

久田敏彦監修(2013)『PISA 後の教育をどうとらえるか—ドイツをとおしてみる—』八千代出版

細尾萌子「フランスの中等教育における基礎学力論争—知識かコンピテンシーか—」『近畿大学教育論叢』第26巻第1号、2014年。

参考・引用文献

Eisner,E.W.,(1985)The Art of Educational Evaluation. The Falmer Press.

藤本奈美(2013)『フレッド・ニューマンの「真正の学力」概念に関する一考察』『教育目標・評価学会紀要』第23号。

二宮衆一(2012)「イギリスのARGによる「学習のための評価」論の考察」『教育方法学研究』第38巻。

天野正輝(1993)『教育評価史研究—教育実践における評価論の系譜』東信堂。

荒井克弘・橋本昭彦編著(2005)『高校と大学の接続 入試選抜から教育接続へ』玉川大学出版部。

石井英真(2010)『現代アメリカにおける学力形成論の展開』東信堂。

稲葉宏雄(1984)『学力問題と到達度評価』上下、あゆみ出版。

ウィギンズ・マクタイ、西岡加名恵訳(2012)『理解をもたらすカリキュラム設計』日本標準。

奥村好美(2012)「オランダにおける『教育の自由』のもとでの学校評価—教育監査とオランダ・ダルトン教会の訪問視察の評価指標に注目して—」『教育方法学研究』第38巻。

遠藤貴広(2004)「ウィギンズの『看破』学習—1980年代後半のエッセンシャル・スクール連盟における『本質的な問い』を踏まえて—」『教育方法学研究』第30巻。

鹿毛雅治(2007)『子どもの姿に学ぶ教師』教育出版。

梶田叡一(1979)『到達度評価の理論と教育革新』明治図書。

梶田叡一(1983)『教育評価』有斐閣。

勝野頼彦(国立教育政策研究所教育課程研究センター(2013)『社会の変化に対応する資質や能力を育成する教育課程編成の基本原理解』。

桂直美(2009)『『ワークショップ授業モデル』による表現の授業構成』『教育方法学研究』第35巻。

川地亜弥子(2004)「生活綴方における教育評価論の構造—1930年代の『集合的合評作業』の分析を中心に—」『教育方法学研究』第30巻。

北川剛司(2006)「コンテンツ・リテラシーとしての書く力の評価に関する研究」『教育方法学研究』第32巻。

京都大学大学院教育学研究科・教育実践コラボレーション・センター(2012)「記録『中韓日シンポジウム アジアにおけるPISA問題』」『活動報告書(2007年度—2011年度)』pp.371-440。

遠山啓等『現代教育科学』1961年3月号所収。

項純(2013)『現代中国における教育評価改革』日本標準。

小見山栄一(1950)「評価の教育的意義」東京教育大学教育学研究室編『教育大学講座32巻 教育評価』金子書房。

佐々木元緒(1979)『到達度評価—その考え方と進め方』明治図書。

佐藤学編著(2009)『揺れる世界の学力マップ』明石書店。

鈴木秀幸(2013)『スタンダード準拠評価』図書文化

Tyler, R. W. (1934). Constructing Achievement Tests. (Ohio State Univ.)

田中耕治(1982)「教育目標とカリキュラム構成の課題—ブルームとアイズナーの所説を中心に—」『京都大学教育学部紀要』第28号。

田中耕治(1995)「OBEの現状と課題」稲葉宏雄編著『教育方法学の再構築』あゆみ出版。

田中耕治(1996)『学力評価論入門』法政出版。

田中耕治(2008a)『教育評価』, 岩波書店。

田中耕治編著(2008b)『新しい学力テストを読み解く』日本標準。

田中耕治(2009)『人物で綴る戦後教育評価の歴史』三学出版。

田中耕治編著(2001)『パフォーマンス評価』ぎょうせい。

田中耕治(2010a)『序 教育目標・評価研究の20年』

「『評価の時代』を読み解く—教育目標・評価研究の課題と展望—」上下, 日本標準。

田中耕治(2010b)『指導要録のあゆみとこれから』田中耕治編著『小学校新指導要録改訂のポイント』日本標準。

田中耕治(2010c)『新しい「評価のあり方」を拓く—「目標に準拠した評価」のこれまでとこれから—』日本標準。

田中耕治(2013)『教育評価と教育実践の課題—「評価の時代」を拓く—』三学出版。

續有恒(1969)『教育評価』第一法規。

續有恒(1973)『教育心理学の探求』金子書房。

坂井誠亮(2009)「初期社会科『田原本プラン』における学習評価の特質」『教育方法学研究』第35巻。

林尚示(1997)「千葉県北条プランにおける能力評価の変遷に関する研究」『カリキュラム研究』第6号。

文部省(1951)『初等教育の原理』東洋館出版。

増田幸一(1955)「エヴァリュエーション(Evaluation)の意義について」『神戸大学教育学部研究集録』。

東井義雄(初出1972)『東井義雄著作集』四巻, 明治図書。

東井義雄(1967)『「通信簿」の改造』明治図書。

滝沢武久(1962)『授業と認識』明治図書。

松本陽一・高橋嘉明編(1988)『斎藤喜博の世界』一荳書房。

續有恒(1973)「第4章 三 評価の意味 四 教育評価の本質」『教育心理学の探求』金子書房, 1973年。

ダイアン・ハート, 田中耕治監訳(2012)『パフォーマンス評価入門—「真正の評価」論からの提案』ミネルヴァ書房。

中内敏夫(1960)「教育評価」『現代教育科学』2, 岩波書店。

中内敏夫(1971)『学力と評価の理論』, 国土社。

中内敏夫(1977)「教育の目標・評価論の課題」『教育』1977年7月号。

中内敏夫(1998)『中内敏夫著作集Ⅰ「教室」をひらく—新・教育原論』藤原書店。

長島貞夫(1949)「エヴァリュエーション」『新教育事典』平凡社。

西岡加名恵(2003)『教科と総合に活かすポートフォリオ評価法』図書文化。

根津朋実(2006)『カリキュラム評価の方法』多賀出版。

橋本重治(1961)「相対評価法と絶対評価法」『教育心理』5月号。

橋本重治(1976)『新・教育評価法総説』金子書房。

久田敏彦監修(2013)『PISA 後の教育をどうとらえるか—ドイツをとおしてみる—』八千代出版。

平田知美(2007)「『発達最近接領域』の評価に関する実践的研究—算数授業におけるダイナミック・アセスメントの試み」『教育方法学研究』第33巻。

平野朝久(1981)「『目標にとらわれない評価(goal-free evaluation)』についての一考察」『教育方法学研究』第32巻。

ブルーム他著、梶田叡一他訳(1973)『教育評価法ハンドブック』第一法規。

ブルーム他著、梶田叡一他訳(1974)『学習評価ハンドブック』上下、第一法規。

ブルーム著、稲葉宏雄・大西匡哉監訳(1986)『すべての子どもにたしかな学力を』明治図書。

細尾萌子(2010)「フランスの教育評価論の歴史的展開」『教育方法学研究』第36巻。

本所恵(2008)「スウェーデンの高校における必修科目の教育目標—数学の全国テストの検討を中心に—」『教育方法学研究』第34巻。

文部省(1951)『初等教育の原理』。

文部省(1975)『カリキュラム開発の課題』。

堀哲夫(2013)『一枚ポートフォリオ評価(OPPA)』東洋館出版社。

正木正(1948)『教育の心理』羽田書店。

正木正(1952)「価値と評価」, 講座『教育』第3巻, 岩波書店, 1952年所収。

村越邦雄(1978)『子どものための教育評価』青木書店。

山本はるか(2013)「米国における言語化スタンダードの分析—『多文化性』を尊重し『結果の平等性』を保障する方との模索」『教育目標・評価学会紀要』第23号。

山下正俊・黒田耕司(1981)「授業過程と評価論の今日的課題(1)」『教育方法学研究』第32巻。

吉崎静夫(1981)「児童による授業評価(2)—評価尺度の妥当性の検討—」『教育方法学研究』第32巻。

吉田成章(2006)「ノイナ(Neuner, G)による東ドイツ教授学の評価に関する研究—『一般陶冶』の構想を中心と

して—」『教育方法学研究』第32巻。

＜この間に出版された注目すべき教育評価に関する文献＞

上田薫(1988)「評価方法の革新」『学力と授業』黎明書房, pp.70-90。(初出 日本教育方法学会編『授業改造の基本問題』明治図書, 1968年所収)。

板倉聖宣(1989)「私の評価論」『私の評価論』国土社, pp.35-66。(初出 『ひと』1974年6月号所収)

長尾彰夫、浜田寿美男編(2000)『教育評価を考える』ミネルヴァ書房。

キャロライン・V・ギブス著、鈴木秀幸訳(2001)『新しい評価を求めて』論創社。

荒井克弘・倉元直樹編著(2008)『全国学力調査 日米比較研究』金子書房。

岩崎保之(2009)『目標準拠評価の研究』ウェストン出版部。

西村和夫他編(2010)『混迷する評価の時代』東信堂。

木原俊行(2011)『活用型学力を育てる授業づくり』ミネルヴァ書房。

永岡慶三編著(2012)『教育工学における学習評価』ミネルヴァ書房。

佐藤浩一編著(2013)『学習支援と教育評価』北大路書房。

マルザーノ、ケンドール著、黒上晴夫、泰山裕訳(2013)『教育目標をデザインする』北大路書房。

☆なお、本稿は日本教育方法学会編『教育方法学研究ハンドブック』学文社, 2014年の所収の拙稿「教育評価」(約6,000字程度)に大幅な加筆修正(約18,000字程度)を行ったものである。